2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT) | 979-8-3503-8354-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/IC2PCT60090.2024.10486458

Kannada Manuscript Digitization through OCR and Machine Learning

Asha Rani Borah
Department of CSE,
New Horizon College of Engineering,
Bengaluru, Karnataka, India
Asha.borah@gmail.com

Abhilash Vijapur
Department of CSE,
New Horizon College of
Engineering, Bengaluru, Karnataka, India
Abhilashv9941@gmail.com

B Mahesh Kumar
Department of CSE,
New Horizon College of Engineering,
Bengaluru, Karnataka, India
Mahesh.bichhali@gmail.com

Abstract- **This research paper deals with the challenges of digitizing ancient Kannada manuscripts, addressing all the complex features of the Kannada script which has different types of handwriting with different styles depending on the historical context. The main focus is on development of a special Optical Character Recognition (OCR) model which can decipher Kannada characters and convert them into digital characters. This process follows CNN algorithm which accepts pre-processed data and predicts each character based on the trained and tested database. Another idea is to predict the approximate age of the manuscript depending on the linguistic features of the Kannada that is written in different eras. Through this, the research can preserve and make old Kannada manuscripts more accessible ensuring future generations can access these manuscripts.**

Keywords— Manuscript, OCR, Kannada Manuscript, CNN.

I. INTRODUCTION

Digitization of the Historical manuscripts is one of the most important things that should be done in order to preserve these for future generations. Kannada language is one of the scripts which is of great importance and has lot of historical manuscripts. Kannada script is one of the complex scripts, which makes it hard to digitize. There can be difference in handwriting, difference in character style and can have multiple contexts based on the era of manuscript.

This research studies the problems and obstacles in maintaining old Kannada manuscripts and proposes a solution which uses OCR and CNN Algorithm for digitizing Kannada Characters.

The Aim of the research is to create a customized system which recognizes Kannada Characters accurately and deals with different styles and also understands context by guessing the age of the manuscript. It uses pre-processing, Algorithm, and post-processing process to digitize language by increasing accuracy.

Additionally, the research includes a prediction angle which predicts the manuscripts age. This is done using different Linguistic features and context of the words. It even considers temporal context.

This paper expands upon the work and advances that have already taken place. The aim is to help in preserving the cultural richness of Kannada by using technology and innovative approach with historical awareness.

II. METHODS AND MATERIAL

A. Preprocessing

1) Characters Preprocessing:

Convert the image into Greyscale for standardizing the color space.

Denoise the image by using denoising techniques and remove extraneous artefacts.

Improve the quality of image by using Autoencoder.

Do Contour detection using OpenCV which can detect accurately.

Crop each characters using Contour crop.

2) Temporal Context Preservation:

Store the date of each inscription for knowing the historical context.

3) Character Recognition

4) Data Preparation

5) Grayscale Image Preprocessing:

Greyscale Image Processing is used for improving the quality of Character recognition. This image is then sent for next step where noise and Sharpness is set for CNN.

6) Noise Reduction and Image Sharpening:

Noise reduction and sharpness techniques are used for increasing the quality and clarity of the image. This is done to increase the signal to noise ratio. This step helps CNN algorithm to train the data.

7) Contour-based Cropping

Contour based cropping is used for cropping characters in the image. Borders are found and each character is cropped independently and this data is fed into CNN algorithm.

8) Data Augmentation

Data Augmentation technique is used generalization skill. This changes dataset by flipping and rotation. our aim is to replicate different writing style and different orientation which are available in Kannada Characters. This increases the accuracy of the digitizing process.

9) Data Segmentation

Data Segmentation is a technique where the image provided is segmented into smaller parts like sentences, words and letters.

Each image is first divided into sentences or lines. These sentences are stored in a folder. This folder contains n number of images, each image consisting of a line.

These images are again segmented into words using contour-based borders. These words are again saved as individual images in a folder.

These are again broken down into individual letters and Ottaksharas and are saved as individual images in a folder. This folder is fed into CNN algorithm which then predicts the letters.

10) Data Slant

After segmentation slant of each letter is corrected so that it will be easier for CNN to predict the letter. It drastically increases the accuracy.

11) Data Thinning

Data thinning is the last step of pre processing where the obtained letters go through thinning process and is send for CNN Code.

B. Mapping

Unicode mapping is done where every letter of dataset is mapped with proper indexing. Sequence dictionary and sequence generation are the files which help in mapping.

C. CNN Model Architecture

1) Design:

To ensure Convoluted Neural Network is properly used for extracting characters from Kannada letters, it is accurately built. Max pooling different layers for downpooling and spatial patterns for collection of convolutional layers is coded. Class probabilities is used for translating extracted characters which is of large architecture.

CNN module contains and uses keras which helps in neural networking. Kannada script is divided into letters and ottaksharas. Ottaksharas are present below the letters for providing a particular sound.

2) Optimization and Loss

Adam optimizer is used with proper learning rate during training which can fine tune our Architecture. Cross entropy is used for loss function for a better match with a accurate character recognition. This is used for loss and optimization.

D. KNN Model Architecture

1) Design

K- nearest neighbors machine learning algorithm is used for prediction numbers. Numbers are better predicted using K-nearest neighbors' algorithm. Similar to letters all the pre-processing is done to numbers and then data is sent to KNN.

2) Optimization and Loss

Similar to CNN Adam optimizer is used with proper learning rate during training which can fine tune our Architecture. Cross entropy is used for loss function for a better match with a accurate character recognition.

E. Training

1) Data Split

For Training, Adam Optimizer is used. Dataset was divided into separate training and testing validation before the training. This gave us a proper baseline for testing which helps us to control the algorithm while training. This data split increases accuracy.

2) Training Methodology

We used enhanced photos for training the model. Accuracy and loss, during this procedure was closely seen, which kept the accuracy high and best possible conversions and performance of the model. With the learning rate was supervised. Cross entropy loss function was used which increases the character recognition accuracy in the application.

F. Validation

1) Test Set

For testing the Data set for validation, we used the different set of data which was designed and it was hidden during the training of the model.

This helped us in evaluation of the model capacity and to find the result on the undisclosed data.

G. Performance review

F1 score, Precision, accuracy and recall, where some of the metrics which was used for review of the performance, Confusion matrices were used for thorough examination, which gives some information about difficulties in the model in identifying the Kannada characters.

H. Fine-Tuning

Fine Tuning of the algorithm was made, which can increase the accuracy, different iterative process was taken into account which increase the efficacy, it consistently increased the character recognition of various historical periods of Kannada.

I. Age Prediction Model

1) Temporal Feature Extraction

2) Feature Selection:

Feature selection is one of the temporal feature Extraction method in age prediction model. Different temporal variations over various historical periods are taken into account, temporal features like script changes and changes in the writing style and different linguistic characteristics that were exclusive to the different eras Kannada literature were used. These were the features that were considered for feature selection.

3) Data Annotation

Data Annotation is the next step where dataset is created with known dates of each inscription. This creates a solid foundation for the training model which can deviate through different eras of Kannada language using the dataset. This is the supervised learning strategy.

J. Regression Model

1) Model Architecture:

RNN, which is Recurrent Neural network model architecture, is used for forecasting the approximate age of the inscription. Our regression model uses temporal dependencies present in our data to predict the age of the manuscript.

2) Training

Training of the model is done with the annotated dataset. The model is trained for predicting the precise age. Model parameters are regularly changed in the training process. Till we See Complex patterns in the data. Cross validation is performed with previously unused data.

K. Evaluation

Mean absolute error (MAE) is used for evaluation which gives us the difference between the estimated and the actual age of the manuscript. Mean absolute error gives us important information about the models accuracy.

L. Iteration and Refinement

Continuous refinement and many iterations are run after evaluation which improves the overall performance of the model and accuracy was gradually increased after some iterations.

M. Web Application

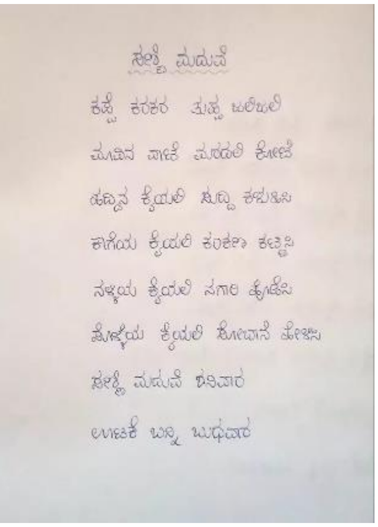
A web Application is created using the Django framework. This web application has two level of access, Admin and User.

Admin access provide list of all users and their details. Admin can add or remove any user.

Users can upload a image and then pre process the image and then see the segmented results and at the end the can download the digital pdf or word document which has the manuscript in typed format.

III. RESULTS

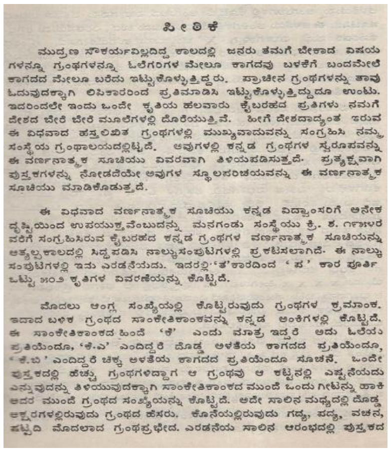
These below images are put into processing



ಹಕ್ಕಿ ಮದುವೆ

ಕಪ್ಪೆ ಕರಕರ ತುಪ್ಪ ಜಲಿಜಲಿ
[illegible] [illegible] ಮರದಲಿ ಕೋಟೆ
ಹದ್ದಿನ ಕೈಯಲಿ [illegible] [illegible]
ಕಾಗೆಯ ಕೈಯಲಿ ಕಂಕಣ ಕಟ್ಟಿಸಿ
ನಳ್ಳಿಯ ಕೈಯಲಿ ನಗಾರಿ ಹೊಡೆಸಿ
[illegible] ಕೈಯಲಿ [illegible] [illegible]
ಹಕ್ಕಿ ಮದುವೆ ಶನಿವಾರ
ಊಟಕೆ ಬನ್ನಿ ಬುಧವಾರ

1)Image before processing one



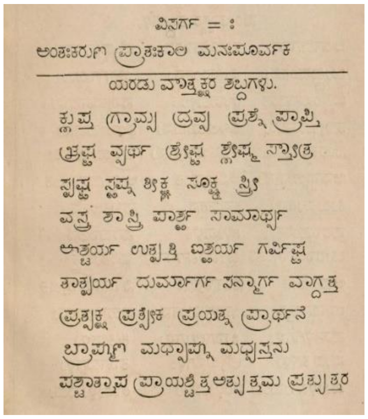
ಪೀಠಿಕೆ

ಮುದ್ರಣ ಸೌಕರ್ಯವಿಲ್ಲದಿದ್ದ ಕಾಲದಲ್ಲಿ ಜನರು ತಮಗೆ ಬೇಕಾದ ವಿಷಯಗಳನ್ನೂ ಗ್ರಂಥಗಳನ್ನೂ ಓಲೆಗರಿಗಳ ಮೇಲೂ ಕಾಗದವು ಬಳಕೆಗೆ ಬಂದಮೇಲೆ ಕಾಗದದ ಮೇಲೂ ಬರೆದು ಇಟ್ಟುಕೊಳ್ಳುತ್ತಿದ್ದರು. ಪ್ರಾಚೀನ ಗ್ರಂಥಗಳನ್ನು ತಾವು ಓದುವುದಕ್ಕಾಗಿ ಲಿಪಿಕಾರರಿಂದ ಪ್ರತಿಮಾಡಿಸಿ ಇಟ್ಟುಕೊಳ್ಳುತ್ತಿದ್ದುದೂ ಉಂಟು. ಇದರಿಂದಲೇ ಇಂದು ಒಂದೇ ಕೃತಿಯ ಹಲವಾರು ಕೈಬರಹದ ಪ್ರತಿಗಳು ನಮಗೆ ದೇಶದ ಬೇರೆ ಬೇರೆ ಮೂಲೆಗಳಲ್ಲಿ ದೊರೆಯುತ್ತವೆ. ಹೀಗೆ ದೇಶದಾದ್ಯಂತ ಇರುವ ಈ ವಿಧವಾದ ಹಸ್ತಲಿಖಿತ ಗ್ರಂಥಗಳಲ್ಲಿ ಮುಖ್ಯವಾದುವನ್ನು ಸಂಗ್ರಹಿಸಿ ನಮ್ಮ ಸಂಸ್ಥೆಯ ಗ್ರಂಥಾಲಯದಲ್ಲಿಟ್ಟಿದೆ. ಅವುಗಳಲ್ಲಿ ಕನ್ನಡ ಗ್ರಂಥಗಳ ಸ್ವರೂಪವನ್ನು ಈ ವರ್ಣನಾತ್ಮಕ ಸೂಚಿಯು ವಿವರವಾಗಿ ತಿಳಿಯಪಡಿಸುತ್ತದೆ. ಪ್ರತ್ಯಕ್ಷವಾಗಿ ಪುಸ್ತಕಗಳನ್ನು ನೋಡದೆಯೇ ಅವುಗಳ ಸ್ಥೂಲಪರಿಚಯವನ್ನು ಈ ವರ್ಣನಾತ್ಮಕ ಸೂಚಿಯು ಮಾಡಿಕೊಡುತ್ತದೆ.

ಈ ವಿಧವಾದ ವರ್ಣನಾತ್ಮಕ ಸೂಚಿಯು ಕನ್ನಡ ವಿದ್ವಾಂಸರಿಗೆ ಅನೇಕ ದೃಷ್ಟಿಯಿಂದ ಉಪಯುಕ್ತವೆಂಬುದನ್ನು ಮನಗಂಡು ಸಂಸ್ಥೆಯು ಕ್ರಿ. ಶ. ೧೯೫೪ರ ವರೆಗೆ ಸಂಗ್ರಹಿಸಿರುವ ಕೈಬರಹದ ಕನ್ನಡ ಗ್ರಂಥಗಳ ವರ್ಣನಾತ್ಮಕ ಸೂಚಿಯನ್ನು ಅತ್ಯಲ್ಪಕಾಲದಲ್ಲಿ ಸಿದ್ಧಪಡಿಸಿ ನಾಲ್ಕುಸಂಪುಟಗಳಲ್ಲಿ ಪ್ರಕಟಿಸಲಾಗಿದೆ. ಈ ನಾಲ್ಕು ಸಂಪುಟಗಳಲ್ಲಿ ಇದು ಎರಡನೆಯದು. ಇದರಲ್ಲಿ 'ತ'ಕಾರದಿಂದ 'ಪ' ಕಾರ ಪೂರ್ತಿ ಒಟ್ಟು ೫೦೨ ಕೃತಿಗಳ ವಿವರಣೆಯನ್ನು ಕೊಟ್ಟಿದೆ.

ಮೊದಲು ಆಂಗ್ಲ ಸಂಖ್ಯೆಯಲ್ಲಿ ಕೊಟ್ಟಿರುವುದು ಗ್ರಂಥಗಳ ಕ್ರಮಾಂಕ. ಇದಾದ ಬಳಿಕ ಗ್ರಂಥದ ಸಾಂಕೇತಿಕಾಂಕವನ್ನು ಕನ್ನಡ ಅಂಕಿಗಳಲ್ಲಿ ಕೊಟ್ಟಿದೆ. ಈ ಸಾಂಕೇತಿಕಾಂಕದ ಹಿಂದೆ 'ಕೆ' ಎಂದು ಮಾತ್ರ ಇದ್ದರೆ ಅದು ಓಲೆಯ ಪ್ರತಿಯೆಂದೂ, 'ಕೆ.ಎ' ಎಂದಿದ್ದರೆ ದೊಡ್ಡ ಅಳತೆಯ ಕಾಗದದ ಪ್ರತಿಯೆಂದೂ, 'ಕೆ.ಬಿ' ಎಂದಿದ್ದರೆ ಚಿಕ್ಕ ಅಳತೆಯ ಕಾಗದದ ಪ್ರತಿಯೆಂದೂ ಸೂಚನೆ. ಒಂದೇ ಪುಸ್ತಕದಲ್ಲಿ ಹೆಚ್ಚು ಗ್ರಂಥಗಳಿದ್ದಾಗ ಆ ಗ್ರಂಥವು ಆ ಕಟ್ಟಿನಲ್ಲಿ ಎಷ್ಟನೆಯದು ಎನ್ನುವುದನ್ನು ತಿಳಿಯುವುದಕ್ಕಾಗಿ ಸಾಂಕೇತಿಕಾಂಕದ ಮುಂದೆ ಒಂದು ಗೀಟನ್ನು ಹಾಕಿ ಅದರ ಮುಂದೆ ಗ್ರಂಥದ ಸಂಖ್ಯೆಯನ್ನು ಕೊಟ್ಟಿದೆ. ಅದೇ ಸಾಲಿನ ಮಧ್ಯದಲ್ಲಿ ದೊಡ್ಡ ಅಕ್ಷರಗಳಲ್ಲಿರುವುದು ಗ್ರಂಥದ ಹೆಸರು. ಕೊನೆಯಲ್ಲಿರುವುದು ಗದ್ಯ, ಪದ್ಯ, ವಚನ, ಷಟ್ಪದಿ ಮೊದಲಾದ ಗ್ರಂಥಪ್ರಭೇದ. ಎರಡನೆಯ ಸಾಲಿನ ಆರಂಭದಲ್ಲಿ ಪುಸ್ತಕದ

2)Image before processing two

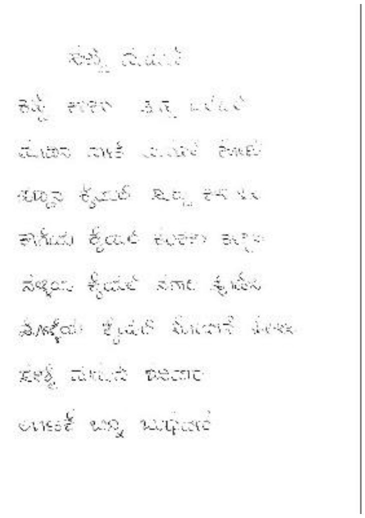


ವಿಸರ್ಗ = ಃ
ಅಂತಃಕರಣ ಪ್ರಾತಃಕಾಲ ಮನಃಪೂರ್ವಕ

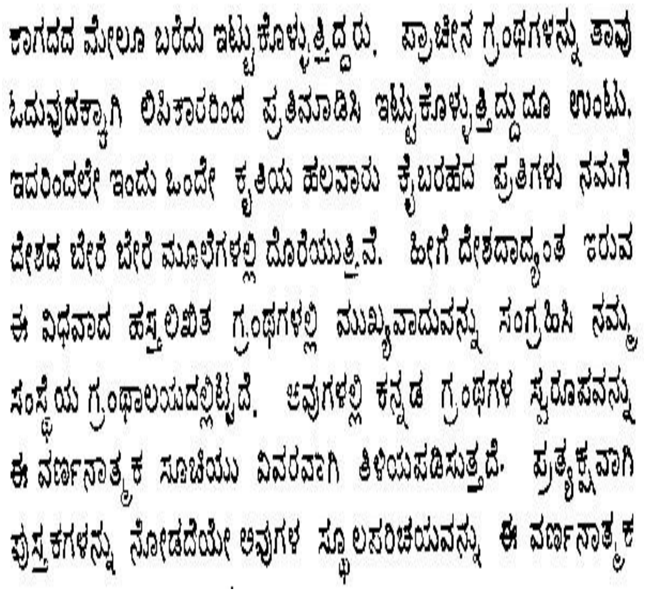
ಯರಡು ಒತ್ತಕ್ಷರ ಶಬ್ದಗಳು.
ಕ್ಲುಪ್ತ ಗ್ರಾಮ್ಯ ದ್ರವ್ಯ ಪ್ರಶ್ನೆ ಪ್ರಾಪ್ತಿ
ಭ್ರಷ್ಟ ವ್ಯರ್ಥ ಶ್ರೇಷ್ಠ ಶ್ಲೇಷ್ಮ ಸ್ತೋತ್ರ
ಸ್ಪಷ್ಟ ಸ್ವಪ್ನ ತೀಕ್ಷ್ಣ ಸೂಕ್ಷ್ಮ ಸ್ತ್ರೀ
ವಸ್ತ್ರ ಶಾಸ್ತ್ರ ಪಾರ್ಶ್ವ ಸಾಮರ್ಥ್ಯ
ಆಶ್ಚರ್ಯ ಉತ್ಪತ್ತಿ ಐಶ್ಚರ್ಯ ಗರಿಷ್ಠ
ತಾತ್ಪರ್ಯ ದುರ್ಮಾರ್ಗ ಸನ್ಮಾರ್ಗ ವಾಗ್ದತ್ತ
ಪ್ರತ್ಯಕ್ಷ ಪ್ರತ್ಯೇಕ ಪ್ರಯತ್ನ ಪ್ರಾರ್ಥನೆ
ಬ್ರಾಹ್ಮಣ ಮಧ್ಯಾಹ್ನ ಮಧ್ಯಸ್ತನು
ಪಶ್ಚಾತ್ತಾಪ ಪ್ರಾಯಶ್ಚಿತ್ತ ಅತ್ಯುತ್ತಮ ಪ್ರತ್ಯುತ್ತರ

3)Image before processing three

Fig. 1. Iimages to be processed



1)Processed image one

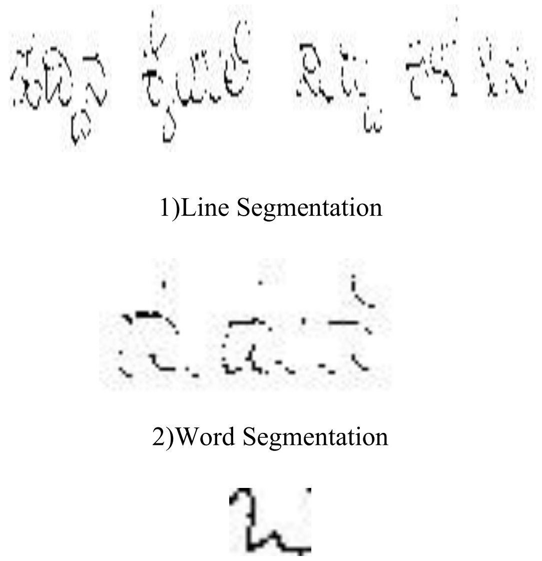


ಕಾಗದದ ಮೇಲೂ ಬರೆದು ಇಟ್ಟುಕೊಳ್ಳುತ್ತಿದ್ದರು. ಪ್ರಾಚೀನ ಗ್ರಂಥಗಳನ್ನು ತಾವು
ಓದುವುದಕ್ಕಾಗಿ ಲಿಪಿಕಾರರಿಂದ ಪ್ರತಿಮಾಡಿಸಿ ಇಟ್ಟುಕೊಳ್ಳುತ್ತಿದ್ದುದೂ ಉಂಟು.
ಇದರಿಂದಲೇ ಇಂದು ಒಂದೇ ಕೃತಿಯ ಹಲವಾರು ಕೈಬರಹದ ಪ್ರತಿಗಳು ನಮಗೆ
ದೇಶದ ಬೇರೆ ಬೇರೆ ಮೂಲೆಗಳಲ್ಲಿ ದೊರೆಯುತ್ತಿವೆ. ಹೀಗೆ ದೇಶದಾದ್ಯಂತ ಇರುವ
ಈ ವಿಧವಾದ ಹಸ್ತಲಿಖಿತ ಗ್ರಂಥಗಳಲ್ಲಿ ಮುಖ್ಯವಾದುವನ್ನು ಸಂಗ್ರಹಿಸಿ ನಮ್ಮ
ಸಂಸ್ಥೆಯ ಗ್ರಂಥಾಲಯದಲ್ಲಿಟ್ಟಿದೆ. ಅವುಗಳಲ್ಲಿ ಕನ್ನಡ ಗ್ರಂಥಗಳ ಸ್ವರೂಪವನ್ನು
ಈ ವರ್ಣನಾತ್ಮಕ ಸೂಚಿಯು ವಿವರವಾಗಿ ತಿಳಿಯಪಡಿಸುತ್ತದೆ. ಪ್ರತ್ಯಕ್ಷವಾಗಿ
ಪುಸ್ತಕಗಳನ್ನು ನೋಡದೆಯೇ ಅವುಗಳ ಸ್ಥೂಲಪರಿಚಯವನ್ನು ಈ ವರ್ಣನಾತ್ಮಕ

2)Processed image two

Fig. 2. Processeed Image

Line word and letter segmentation of image one



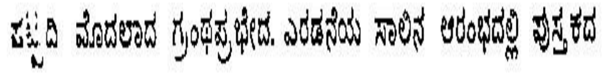
1)Line Segmentation

2)Word Segmentation

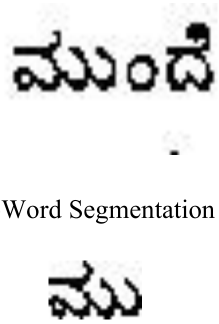
3)Letter Segmentation

Fig. 3. Line word and letter segmentation of image one

Line word and letter segmentation of image two


ಷಟ್ಪದಿ ಮೊದಲಾದ ಗ್ರಂಥಪ್ರಭೇದ. ಎರಡನೆಯ ಸಾಲಿನ ಆರಂಭದಲ್ಲಿ ಪುಸ್ತಕದ

1)Line Segmentation

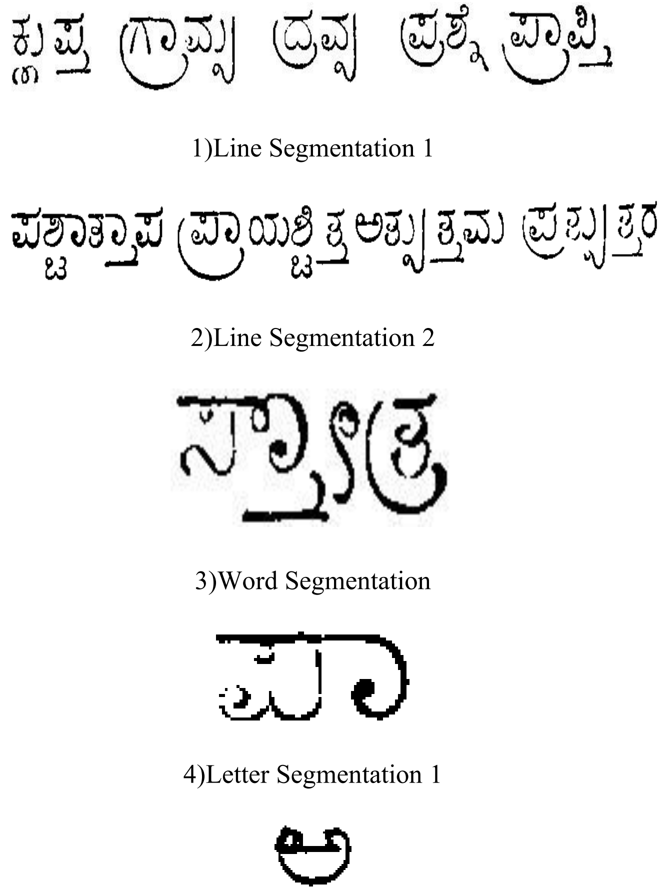


2)Word Segmentation

3)Letter Segmentation

Fig. 4. Line word and letter segmentation of image two

Line word and letter segmentation of image three


ಕ್ಲೃಪ್ತ ಗ್ರಾಮ್ಯ ದ್ರವ್ಯ ಪ್ರಶ್ನೆ ಪ್ರಾಪ್ತಿ
ಪಶ್ಚಾತ್ತಾಪ ಪ್ರಾಯಶ್ಚಿತ್ತ ಅಪ್ಪುತ್ತಮ ಪ್ರಪ್ಪುತ್ತರ

1)Line Segmentation 1

2)Line Segmentation 2

3)Word Segmentation

4)Letter Segmentation 1

5)Letter Segmentation 2

Fig. 5. Line word and letter segmentation of image three

This Data is fed into CNN, KNN Algorithms which provide the digital copy. Accuracy is 68 percent.

Copyright © IEEE–2024 ISBN: 979-8-3503-8354-6

IV. APPLICATIONS

Optical character recognition of the old manuscripts has various applications and it promises variety of uses.

A. Cultural Preservation

Cultural preservation is one of the application of digitization of old Kannada manuscripts as upcoming generations can have access to old Kannada manuscripts and they can learn Kannada language, its history and other customs of different era of Kannada language.

B. Educational Resources

Linguistic researchers and scholars can access this digital manuscripts which are one of the invaluable resources for historical and cultural studies. The materials and different manuscripts can be easily accessible.

C. Historical Research

Historical research is one of the application of digitization of Kannada language as researchers can examine and analyse the alterations that have taken place in Kannada from the old times.

D. Museum Exhibits and Archives

Museum Exhibits and archives can display old Kannada manuscripts which can be open for public to learn and research about Kannada language.

V. CONCLUSION

To conclude this paper deals with the Complex problems of digitization of Kannada language and it uses OCR and CNN for digitization. Through this digitization, it preserves old manuscripts, which can be of cultural importance.

The main goal of this research is to create a machine learning model which can digitize Kannada characters from the images taken of different manuscripts.

Traditional way of character recognition is improved by using post processing techniques and sophisticated segmentation strategies in preprocessing. These advanced help in accuracy of this digitization.

Age prediction is also included, which is one of the important feature where time dimension is considered and historical context is also considered for predicting the accurate age of the manuscript based on the linguistic features.

This research paper proposes a fusion model where historical data is used along with machine learning techniques. This closes the gap between past and the present by using CNN and RNN models. This ensures preserving old Kannada manuscripts and the Legacy of Kannada history.

Three step process of preprocessing, CNN algorithm and post processing has increased the accuracy by attaining the optimal performance And age prediction is also introduced.

As the study and research moves further, it supports the overall goal of preserving old Kannada manuscripts. Further improvements can be made in accuracy of the digitization process.

REFERENCES

[1] Discrete Artificial Bee Colony Algorithm based Optical Character Recognition Nishal Ancelette Pereira, Prajwal Rao, Akshay K Kallianpur and K G Srinivasa, IEEE

[2] Era Identification and Recognition of Ganga and Hoysala Phase Kannada Stone Inscriptions Characters using Advance Recognition Algorithm Dr. H S Mohana, Mr. Rajithkumar B K, IEEE

[3] Read and Recognition of old Kannada Stone Inscriptions Characters using Novel Algorithm Rajithkumar B K, Dr. H.S. Mohana, Uday J, Bhavana M B and Anusha L S, IEEE

[4] Kannada Handwritten Script Recognition using Machine Learning Techniques Roshan Fernandes and Anisha P Rodrigues, IEEE

[5] Comparative Analysis of Algorithms for Recognizing Emotions by Eye Blink Borah, A.R., Subhashini, S.J., Mohesh, A., Roshini, K. 2022 International Conference for Advancement in Technology, ICONAT 2022, 2022

[6] Analysis of Trending NFTs using Time-Series Techniques Borah, A.R., Kumar, S., Kasish, S.V., Jashwanth, M.S., Uma Shankar Reddy, M. 2022 International Conference on Smart Generation Computing, Communication and Networking, SMART GENCON 2022, 2022

[7] Eye Blink—A Mode of Communication Singh, H., Borah, A.R., Suthar, N.H., Reddy, V.R., Ashok, K. Lecture Notes in Networks and Systems

[8] Handwritten Character Segmentation for Kannada Scripts C. Naveena V.N. Manjunath Aradhya, IEEE.

[9] Kannada Text Line Extraction Based on Energy Minimization and Skew Correction Sunanda Dixit, Suresh Hosahalli Narayan, Mahesh Belur, IEEE

[10] Recognition of Ancient Kannada Epigraphs using Fuzzy-Based Approach Soumya A and G Hemantha Kumar, IEEE

[11] Interactive Segmentation for Character Extraction In Stone Inscriptions H.S. Mohana, Navya. K, Rajithkumar. B.K and Nagesh. C, IEEE

[12] ”Classification of Ancient Epigraphs into different periods using Random Forests”, Soumya A and G Hemantha Kumar, IEEE

[13] “Degraded character recognition from old Kannada documents”, Sridevi Tumkur Narasimhaiah, Lalitha Rangarajan, IEEE

Copyright © IEEE–2024 ISBN: 979-8-3503-8354-6